Multi-wavelength anomalous diffraction at high x-ray intensity

Sang-Kil Son (손상길),[1, *] Henry N. Chapman,[1, 2, †] and Robin Santra[1, 2, ‡]

[1]*Center for Free-Electron Laser Science, DESY, Hamburg, Germany*
[2]*Department of Physics, University of Hamburg, Hamburg, Germany*
(Dated: October 18, 2011)

The multi-wavelength anomalous diffraction (MAD) method is used to determine phase information in x-ray crystallography by employing dispersion corrections from heavy atoms on coherent x-ray scattering. X-ray free-electron lasers (FELs) show promise for revealing the structure of single molecules or nanocrystals within femtoseconds, but the phase problem remains largely unsolved. Due to the ultrabrightness of x-ray FEL, samples experience severe electronic radiation damage, especially to heavy atoms, which hinders direct implementation of the MAD method with x-ray FELs. We propose a generalized version of the MAD phasing method at high x-ray intensity. We demonstrate the existence of a Karle–Hendrickson-type equation for the MAD method in the high-intensity regime and calculate relevant coefficients with detailed electronic damage dynamics of heavy atoms. Our results show that the bleaching effect on the scattering strength of the heavy atoms can be advantageous to the phasing method. The present method offers a potential for *ab initio* structural determination in femtosecond x-ray nanocrystallography.

PACS numbers: 87.53.−j, 61.46.Hk, 41.60.Cr, 32.90.+a

Determination of the 3D structure of proteins and macromolecules is crucial to understand their biological functions at the molecular level. X-ray crystallography has been widely used for structural determination [1], but it suffers from two bottlenecks: the phase problem and growing high-quality crystals. The phase problem [2, 3] is a fundamental obstacle in constructing an electronic density map from x-ray diffraction. Multi-wavelength anomalous diffraction (MAD) [4–6] with synchrotron radiation is one of the major achievements to address this issue. X-ray free-electron lasers (FELs) [7] promise to have a revolutionary impact on molecular imaging [8, 9], overcoming the crystal bottleneck. The unprecedented high x-ray fluence provides a sufficiently large number of photons to enable structure determination from diffraction measurements of streams of single molecules [9–11] and nanocrystals [12, 13]. However, due to an extremely high fluence that is ∼100 times larger than the conventional damage limit [14], samples are subject to severe radiation damage [15]. The ultrashort x-ray pulses generated by x-ray FELs enable us to carry out "diffraction-before-destruction" within femtosecond timescales to suppress nuclear motion [9]. Nonetheless, electronic damage [9, 16, 17] during femtosecond x-ray pulses is unavoidable, leading us to consider "diffraction-during-ionization" [18]. This electronic radiation damage is particularly challenging when addressing the phase problem by anomalous dispersion [19], because heavy atoms as anomalous scatterers will be more ionized than other atoms during intense x-ray pulses. Therefore, it has been speculated that MAD would not be an applicable route for phasing in the presence of severe radiation damage [1, 19]. Here we propose a high-intensity version of the MAD phasing method based on a detailed description of the electronic response at the atomic level. In contrast to the speculation, our results show that MAD not only works, but also that the extensive electronic rearrangements at high x-ray intensity provide a new path to phasing. We will then demonstrate that this approach is appli-

cable to the phase problem in femtosecond x-ray nanocrystallography [12, 13], which is one of the most prominent topics in x-ray FEL applications.

X rays mainly ionize inner-shell electrons and subsequent relaxation (Auger decay and fluorescence) fills the inner-shell vacancy. Therefore, sequences of photoionization and relaxation can strip off many electrons after absorbing several photons [20]. For heavy atoms that have more than two subshells, a vacancy in deep inner-shells causes several relaxation steps in the cascade through the subshells, resulting in further electron ejections [21]. To simulate the electronic damage dynamics, we use the XATOM toolkit [18, 22], where all rates and cross sections are calculated within the nonrelativistic Hartree–Fock–Slater method and multiphoton electronic dynamics is described by sequential one-photon processes with all possible electronic configurations. Figure 1 depicts the time evolution of populations for several charge states of an iron (Fe) atom where 27,783 coupled rate equations were solved. The photon energy is 8 keV and the fluence is 5×10^{12} photons/μm². The pulse duration is 10 fs full-width-at-half-maximum (FWHM) with a Gaussian envelope. In these conditions, the neutral Fe is completely depleted and high charge states such as Fe^{20+} are substantially produced by the end of the pulse. The pulse-weighted charge state averaged over time is about +12, demonstrating the severe electronic damage incurred during the x-ray pulse.

This electronic damage affects not only the coherent scattering atomic form factor but also its dispersion correction. Near an inner-shell absorption edge, resonant elastic scattering causes the atomic form factor to depend on the photon energy ω,

$$f(\mathbf{Q}, \omega) = f^0(\mathbf{Q}) + f'(\omega) + if''(\omega), \qquad (1)$$

where $\mathbf{Q}$ is the photon momentum transfer. The XATOM toolkit has been extended to compute the dispersion correction, $f' + if''$ [22]. In Fig. 2, one can see remarkable changes
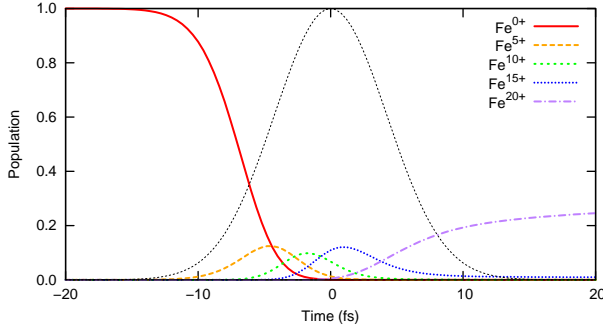
FIG. 1. (Color online) Population dynamics for several selected charge states of Fe during an x-ray pulse. The thin dotted line indicates the Gaussian pulse envelope. See the text for parameters used.
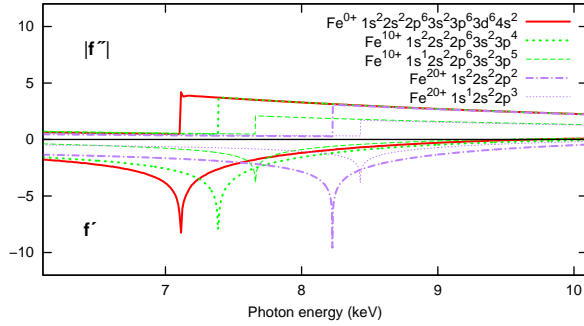

FIG. 2. (Color online) Dispersion corrections of atomic form factors for selected configurations of several charge states of Fe.

of the dispersion correction for different configurations and different charge states of Fe. Both f' and f'' have a singular position at the K-shell edge, which is shifted to a higher ω by ~ 1 keV as the charge state increases. The plotted curves in Fig. 2 correspond to the configurations of the ground state and the single-core-hole state (except for the neutral Fe) for given charge states. Since the MAD phasing method is based on the dispersion correction of heavy elements, it is inevitably required to take into account the electronic damage dynamics and accompanying changes of the dispersion correction under intense x-ray pulses.

In the MAD phasing method, the Karle–Hendrickson equation [23, 24] represents a set of equations of scattering cross sections at several different wavelengths (photon energies). The molecular scattering form factor is separated into normal and anomalous scattering terms and the phase information can be derived from their interferences. In this Letter, we show that a Karle–Hendrickson-type equation exists in the high-intensity regime with extensive electronic damage on anomalous scatterers.

Let P be any protein (or any macromolecule) whose structure we want to solve by coherent x-ray scattering. Let H indicate heavy atoms and N_H be the number of heavy atoms per macromolecule to be considered. Note that P excludes H. Our assumption is that only heavy atoms scatter anomalously

and undergo damage dynamics during an x-ray pulse. It is justified by the fact that the photon energy of interest is near the inner-shell ionization threshold of heavy atoms and the photoabsorption cross section σ of the heavy atom is much higher than that of the light atom for a given range of ω. For example, $\sigma_{Fe}/\sigma_C \approx 300$ at 8 keV and there is almost no dispersion effect on carbon (C) near this photon energy (see Fig. S1(b) in Ref. [25]). The scattering intensity (per unit solid angle) is evaluated by [25],

$$\frac{dI(\mathbf{Q},\omega)}{d\Omega} = \mathscr{F}C(\Omega)\int_{-\infty}^{\infty}dt\,g(t)\sum_I P_I(t)$$
$$\times\left|F_P^0(\mathbf{Q}) + \sum_{j=1}^{N_H}f_{I_j}(\mathbf{Q},\omega)e^{i\mathbf{Q}\cdot\mathbf{R}_j}\right|^2, \qquad (2)$$

where j denotes a heavy atom index and I indicates a global configuration index. The global configuration for N_H heavy atoms is given by $I = (I_1, I_2, \cdots, I_{N_H})$. Here I_j indicates the electronic configuration of the j-th heavy atom, which is located at position $\mathbf{R}_j$. This electronic configuration provides, among other things, information on the charge state of the atom. $P_I(t)$ is the population of the I-th configuration at time t. It is assumed that the heavy atoms are ionized independently, so the population of I is given by a product of individual populations, $P_I(t) = \Pi_{j=1}^{N_H}P_{I_j}(t)$. $\mathscr{F}$ is the x-ray fluence and $g(t)$ is the normalized pulse envelope. Then the x-ray flux is given by $\mathscr{F}g(t)$, which is assumed to be spatially uniform throughout the sample. $C(\Omega)$ is a coefficient given by the polarization of the x-ray pulse.

In Eq. (2), $F_P^0(\mathbf{Q})$ is the molecular form factor for the protein (without any dispersion correction) and our purpose is to solve its amplitude and phase, $F_P^0(\mathbf{Q}) = |F_P^0(\mathbf{Q})|\exp[i\phi_P^0(\mathbf{Q})]$. $f_{I_j}(\mathbf{Q},\omega)$ is the atomic form factor (with the dispersion correction) of the j-th heavy atom in its I_j-th configuration. It is most instructive to consider only one heavy atomic species. We introduce a molecular form factor for undamaged heavy atoms,

$$F_H^0(\mathbf{Q}) = |F_H^0(\mathbf{Q})|e^{i\phi_H^0(\mathbf{Q})} = f_H^0(\mathbf{Q})\sum_{j=1}^{N_H}e^{i\mathbf{Q}\cdot\mathbf{R}_j}, \qquad (3)$$

where $f_H^0(\mathbf{Q})$ indicates the normal scattering atomic form factor for the ground-state configuration of the neutral heavy atom.

Now Eq. (2) can be expanded to demonstrate the existence of a generalized Karle–Hendrickson equation [25],

$$\frac{dI(\mathbf{Q},\omega)}{d\Omega} = \mathscr{F}C(\Omega)\Big[\left|F_P^0(\mathbf{Q})\right|^2 + \left|F_H^0(\mathbf{Q})\right|^2\tilde{a}(\mathbf{Q},\omega)$$
$$+ \left|F_P^0(\mathbf{Q})\right|\left|F_H^0(\mathbf{Q})\right|b(\mathbf{Q},\omega)\cos\left(\phi_P^0(\mathbf{Q}) - \phi_H^0(\mathbf{Q})\right)$$
$$+ \left|F_P^0(\mathbf{Q})\right|\left|F_H^0(\mathbf{Q})\right|c(\mathbf{Q},\omega)\sin\left(\phi_P^0(\mathbf{Q}) - \phi_H^0(\mathbf{Q})\right)$$
$$+ N_H\left|f_H^0(\mathbf{Q})\right|^2\left\{a(\mathbf{Q},\omega) - \tilde{a}(\mathbf{Q},\omega)\right\}\Big], \qquad (4)$$

where the new MAD coefficients depending on $\mathbf{Q}$ and ω are

defined by

$$a(\mathbf{Q},\omega) = \frac{1}{\left\{f_H^0(\mathbf{Q})\right\}^2}\sum_{I_H}\bar{P}_{I_H}\left|f_{I_H}(\mathbf{Q},\omega)\right|^2, \tag{5a}$$

$$b(\mathbf{Q},\omega) = \frac{2}{f_H^0(\mathbf{Q})}\sum_{I_H}\bar{P}_{I_H}\left\{f_{I_H}^0(\mathbf{Q})+f_{I_H}'(\omega)\right\}, \tag{5b}$$

$$c(\mathbf{Q},\omega) = \frac{2}{f_H^0(\mathbf{Q})}\sum_{I_H}\bar{P}_{I_H}f_{I_H}''(\omega), \tag{5c}$$

$$\tilde{a}(\mathbf{Q},\omega) = \frac{1}{\left\{f_H^0(\mathbf{Q})\right\}^2}\int_{-\infty}^{\infty}dt\,g(t)\left|\tilde{f}_H(\mathbf{Q},\omega,t)\right|^2. \tag{5d}$$

Here I_H indicates the electronic configuration of the heavy atom species and $\bar{P}_{I_H} = \int_{-\infty}^{\infty}dt\,g(t)P_{I_H}(t)$ is the pulse-weighted averaged population for the I_H-th configuration. The new MAD coefficients from Eq. (5a) to Eq. (5d) are atom-specific and must be calculated with electronic damage dynamics and configuration-specific atomic form factors. The coefficient a is an incoherent average of $|f_{I_H}|^2$ with $\bar{P}_{I_H}$. The coefficients b and c are the real and imaginary components of the averaged atomic form factor, respectively. The coefficient $\tilde{a}$ in Eq. (5d) is obtained through a *dynamical* form factor defined by

$$\tilde{f}_H(\mathbf{Q},\omega,t) = \sum_{I_H}P_{I_H}(t)f_{I_H}(\mathbf{Q},\omega), \tag{6}$$

which is a coherent average of the configuration-specific form factors over I_H at a given time t. This $\tilde{a}$ coefficient thus represents the effective scattering strength of the heavy atom. In contrast to the original Karle–Hendrickson equation, Eq. (4) is separated into light atoms (P) and heavy atoms (H) because both electronic damage and anomalous scattering are treated exclusively on H. If only the ground-state configuration is considered, i.e., no electronic damage occurs, then $a = \tilde{a}$ and Eqs. (4) and (5) are reduced to the original Karle–Hendrickson equation except for the separation of P and H.

This generalized Karle–Hendrickson equation constitutes a set of equations with different ω at every $\mathbf{Q}$. In Eq. (4) there are three unknowns: $\left|F_P^0(\mathbf{Q})\right|$, $\left|F_H^0(\mathbf{Q})\right|$, and $\phi_P^0(\mathbf{Q})-\phi_H^0(\mathbf{Q})$ for a given $\mathbf{Q}$. With three or more different ω, those unknowns can be solved by the least-square method [26, 27]. Combined with Patterson or direct methods [28, 29], the amplitude and phase of heavy atoms can be determined, so two unknowns of $\left|F_P^0(\mathbf{Q})\right|$ and $\phi_P^0(\mathbf{Q})$ are to be solved with two different ω. Once all amplitudes and phases of P and H are determined, it is straightforward to construct the total structure of $T = P + H$. To obtain non-trivial solutions from the least-square method, the contrast between the coefficients at two different ω must be non-zero. This condition is fulfilled even in the presence of severe electronic damage as shown in the following discussion. We emphasize that, if the MAD coefficients are predetermined experimentally or theoretically, then one can solve the structure (amplitude and phase) from diffraction measurements directly, without any iterative phase retrieval algorithms [30].

Let us consider Fe atoms embedded in a protein and then radiate an x-ray pulse of 2×10^{12} photons and 10 fs FWHM
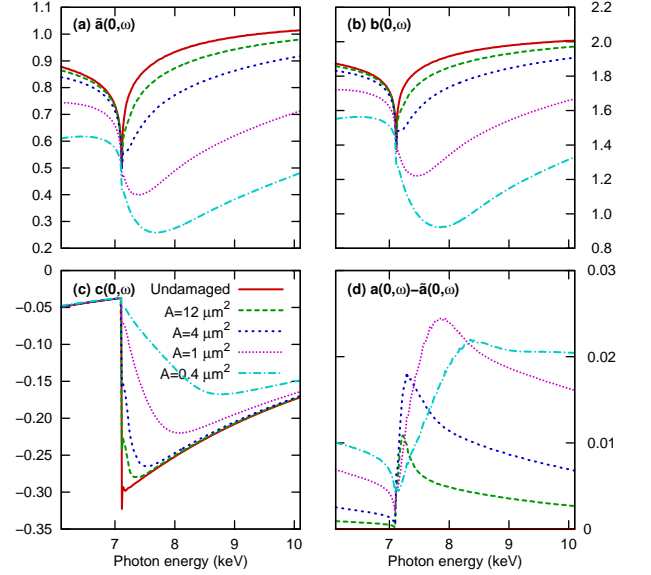

FIG. 3. (Color online) Coefficients in the generalized Karle–Hendrickson equation for Fe as a function of the photon energy. The fluence is given by 2×10^{12} photons/A where A is the focal spot area.

into the sample. Figure 3 displays $\tilde{a}$, b, c, and $(a-\tilde{a})$ for the forward direction ($\mathbf{Q}=\mathbf{0}$) computed by the extended XATOM toolkit [22]. The fluence $\mathscr{F}$ is given by 2×10^{12} photons/A where A is the focal spot area. When high charge states are generated by ionization dynamics, the scattering strength is lowered due to the reduced number of scattering electrons and the change of the dispersion correction. The degree of lowering in $\tilde{a}$ and b shows different behaviors below and above the neutral Fe edge. Below the edge, the scattering strength is less lowered than above the edge because ionization dynamics are dominantly initiated by L-shell ionization whose cross section is 8 times smaller than that of K-shell ionization [31]. Above the edge, K-shell ionization channels are open and lead to further cascade decays, stripping off more electrons. As a result, $\tilde{a}$ and b are dramatically bleached out and their minimum is deepened and broadened. The absolute value of c, which corresponds to the averaged absorption cross section, is decreased as the fluence increases.

This bleaching effect on the scattering strength is beneficial to the phasing problem in two ways. First, the contrast of the coefficients to be exploited in the MAD method is enhanced. Even though the scattering strength is lowered for all ω, Figs. 3(a)–(c) clearly show that the low intensity cases (long-dashed and short-dashed lines) display a contrast similar to the conventional MAD method (solid lines). For the high intensity cases (dotted and dash-dotted lines), the contrast in $\tilde{a}$ and b becomes even larger when ω is chosen below the edge and around the minimum. The contrast in c is reduced to some extent but is not completely eliminated. It is worthwhile to note that broadening of the edge at high intensity makes precision of ω less important in experiments. Second, it brings an alternative phasing method similar to

single isomorphic replacement (SIR) [3] or radiation-damage induced phasing (RIP) [32]. By choosing one ω below and one ω above the edge, one can create two datasets that differ only in the scattering strength of the heavy atoms, and then solve the rest of the structure by density modification. In this method, there is neither atomic replacement in sample preparation like SIR nor chemical rearrangement during the x-ray pulses like RIP. Therefore, the rest of the structure remains invariant in the two different datasets.

Now we discuss experimental implementation of the generalized version of the MAD phasing method. We used a Gaussian pulse in the above calculations. However, when the scattering strength of $dI/d\Omega$ is measured at a particular $\mathbf{Q}$ and ω, the pulse shape $g(t)$ and the fluence $\mathscr{F}$ at a given position $\mathbf{x}$ in the x-ray beam may differ from shot to shot. Following the procedure in Ref. [33], we numerically confirmed that $\langle dI[\mathscr{F}(\mathbf{x})g(t)]/d\Omega \approx dI[\langle \mathscr{F}(\mathbf{x})g(t)\rangle]/d\Omega$ to within 3%, where $\langle \cdot \rangle$ denotes an ensemble average. Then the total signal can be obtained by integrating over the interaction volume, $\int d^3x\, dI\left[\langle \mathscr{F}(\mathbf{x})g(t)\rangle\right]/d\Omega \times n_{\mathrm{mol}}(\mathbf{x})$, where $n_{\mathrm{mol}}(\mathbf{x})$ is the molecular number density. In this process, the MAD coefficients are calculated with given $\langle \mathscr{F}g(t)\rangle$, and the basic structure of Eq. (4) remains unchanged.

In our model, resonant absorption processes and shake-up/shake-off processes [34] are neglected. They tend to generate further high charge states, so the contrast enhancement and the bleaching effect would remain after inclusion of these processes. We note that the effect of impact ionization [35] on coherent diffractive imaging may be suppressed by using a sufficiently short x-ray pulse [18].

The structure of the generalized Karle–Hendrickson equation [Eq. (4)] can be fully functional for phasing of nanocrystals, which are of current interest for structural determination with x-ray FELs [12, 13]. In Eq. (3), F_H^0 contains the structure factors of the heavy atoms. In the case of crystals, the heavy atoms are regularly located and can contribute to the Bragg peaks when satisfying $\mathbf{Q}\cdot(\mathbf{R}_i - \mathbf{R}_j) = 2\pi n$ (n: integer) for all i and j. In Eq. (5d), $\tilde{a}$ is expressed with the coherent average over configurations, and $|F_H^0|^2$ from Eq. (3) is expressed with the coherent summation over heavy atoms. Therefore, $|F_H^0|^2\tilde{a} = \int_{-\infty}^{\infty} dt\, g(t)\left|\tilde{f}_H(\mathbf{Q},\omega,t)\sum_{j=1}^{N_H}\exp[i\mathbf{Q}\cdot\mathbf{R}_j]\right|^2$ implies that all heavy atoms are described by the same *dynamical* form factor. This term is then responsible for the Bragg peaks ($\propto N_H^2$). On the other hand, the term $N_H|f_H^0|^2(a-\tilde{a})$ represents fluctuations from all different configurations induced by electronic damage dynamics, corresponding to the diffuse background ($\propto N_H$). As shown in Fig. 3(d), $(a-\tilde{a})$ increases as the fluence increases. However, it is an order of magnitude smaller than $\tilde{a}$ and not confined to the Bragg peaks, implying that the high x-ray intensity does not fully destroy the coherent signals.

In conclusion, we have proposed the MAD phasing method in extreme conditions of ionizing x-ray radiations. We assume that the scattering factors of the light atoms of the protein do not vary significantly over the measured range of x-ray fre-

quencies, and that these atoms have normal scattering and no ionization. It is also assumed that the heavy atoms are ionized independently and only one type of heavy atoms is considered. We believe that the method should work even if these assumptions are removed, because the most important consequence of high-intensity x-ray irradiation — multiple ionization of the heavy atomic species — has been fully taken into account. We have combined electronic response at the atomic level and molecular imaging during intense x-ray pulses, and demonstrated the existence of a generalized Karle–Hendrickson equation for the MAD method at high x-ray intensity. The relevant coefficients to be used in the MAD method have been formulated and calculated with damage dynamics and accompanying changes of the dispersion correction. We have shown that the generalized equation is still applicable to the phase problem even in the presence of severe radiation damage. The bleaching effect on the scattering strength of heavy atoms, which unexpectedly enhances the coefficient contrast in the MAD method, can be beneficial to phasing. Our study opens up a new opportunity of solving the phase problem in femtosecond nanocrystallography with x-ray FELs.

* sangkil.son@cfel.de
† henry.chapman@cfel.de
‡ robin.santra@cfel.de

[1] Z. Dauter, Acta Cryst. **D62**, 1 (2006).
[2] J. Karle and H. Hauptman, Acta Cryst. **3**, 181 (1950).
[3] G. Taylor, Acta Cryst. **D59**, 1881 (2003).
[4] J. M. Guss *et al.*, Science **241**, 806 (1988).
[5] W. A. Hendrickson *et al.*, Proc. Natl. Acad. Sci. U. S. A. **86**, 2190 (1989).
[6] W. A. Hendrickson, Science **254**, 51 (1991).
[7] B. W. J. McNeil and N. R. Thompson, Nature Photon. **4**, 814 (2010).
[8] K. J. Gaffney and H. N. Chapman, Science **316**, 1444 (2007).
[9] R. Neutze *et al.*, Nature **406**, 752 (2000).
[10] H. N. Chapman *et al.*, Nature Phys. **2**, 839 (2006).
[11] H. N. Chapman and K. A. Nugent, Nature Photon. **4**, 833 (2010).
[12] A. P. Mancuso *et al.*, Phys. Rev. Lett. **102**, 035502 (2009).
[13] H. N. Chapman *et al.*, Nature **470**, 73 (2011).
[14] R. Henderson, Q. Rev. Biophys. **28**, 171 (1995).
[15] M. R. Howells *et al.*, J. Electron Spectrosc. Relat. Phenom. **170**, 4 (2009).
[16] S. P. Hau-Riege, R. A. London, and A. Szoke, Phys. Rev. E **69**, 051906 (2004).
[17] H. M. Quiney and K. A. Nugent, Nature Phys. **7**, 142 (2011).
[18] S.-K. Son, L. Young, and R. Santra, Phys. Rev. A **83**, 033402 (2011).
[19] R. B. G. Ravelli *et al.*, J. Synchrotron Radiat. **12**, 276 (2005).
[20] L. Young *et al.*, Nature **466**, 56 (2010).
[21] T. A. Carlson, W. E. Hunt, and M. O. Krause, Phys. Rev. **151**, 41 (1966).
[22] S.-K. Son and R. Santra, XATOM — an integrated toolkit for

x-ray and atomic physics, CFEL, DESY, Hamburg, Germany, 2011, Rev. 398.

[23] J. Karle, Int. J. Quant. Chem. Quant. Bio. Symp. **7**, 357 (1980).

[24] W. A. Hendrickson, Trans. Am. Crystalgr. Assoc. **21**, 11 (1985).

[25] See Supplemental Material at [URL will be inserted by publisher] for a detailed derivation.

[26] J. Karle, Acta Cryst. **A45**, 303 (1989).

[27] W. A. Hendrickson *et al.*, Proteins: Struct., Funct., Bioinf. **4**, 77 (1988).

[28] H. Hauptman, Angew. Chem., Int. Ed. **25**, 603 (1986).

[29] J. Karle, Science **232**, 837 (1986).

[30] R. P. Millane, J. Opt. Soc. Am. A **7**, 394 (1990).

[31] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Dover, Mineola, NY, 2008).

[32] R. B. Ravelli *et al.*, Structure **11**, 217 (2003).

[33] N. Rohringer and R. Santra, Phys. Rev. A **76**, 033416 (2007).

[34] P. Persson *et al.*, Protein Sci. **10**, 2480 (2001).

[35] B. Ziaja, R. A. London, and J. Hajdu, J. Appl. Phys. **97**, 064905 (2005).